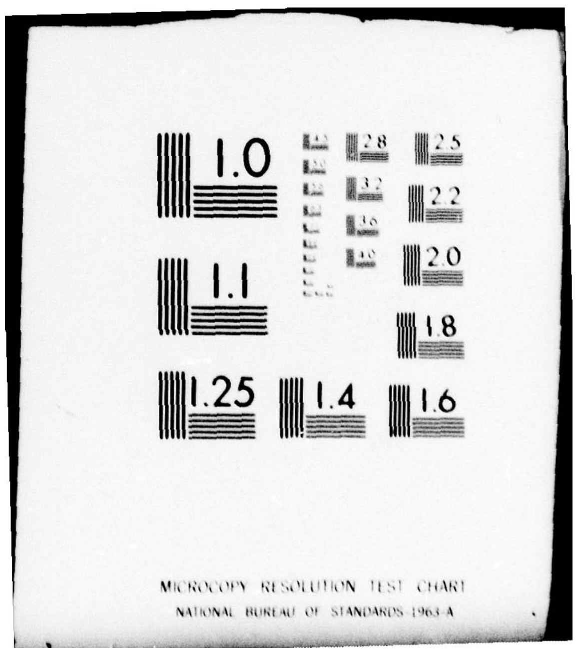
AD-A073 624 NORTH CAROLINA UNIV AT CHAPEL HILL F/G 12/1
A STOPPING RULE FOR DATA COLLECTION IN QUEUEING SIMULATIONS.(U)
AUG 79 V G ADLAKHA, G S FISHMAN N00014-76-C-0302
UNCLASSIFIED TR-79-4 NL

1 OF 1
AD
A073624

END
DATE
FILMED
10-79
DDC


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

13 LEVEL II

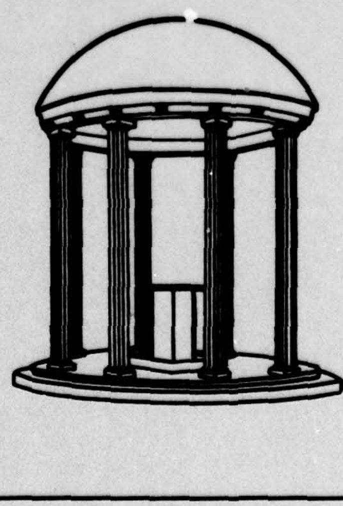
AD A073624

OPERATIONS RESEARCH AND SYSTEMS ANALYSIS

UNIVERSITY OF NORTH CAROLINA
AT CHAPEL HILL

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

DDC FILE COPY



DDC
RECEIVED
SEP 11 1979
B

79 09 10 025

(13) LEVEL II

A STOPPING RULE FOR DATA COLLECTION IN QUEUEING SIMULATIONS.

Veena G. Adlakha ~~and~~ George S. Fishman

Technical Report No. 79-4
August, 1979

33p.

TR-79-4

DDC
RECEIVED
SEP 11 1979
B

Curriculum in Operations Research
and Systems Analysis

University of North Carolina at Chapel Hill

This research was supported by the Office of Naval Research under contract N00014-76-C-0302.

Reproduction in whole or in part is permitted for any purpose of the United States government.

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

259 500

-A-

ABSTRACT

This paper proposes a stopping rule (Rule 2) for terminating data collection in a queueing simulation experiment. The appeal of the rule lies in the fact that data collected in this way can be used to compute interval estimates with coverage rates that compare favorably with theoretically specified rates. The rule relies on a comparison between *a priori* information on the activity level (traffic intensity) ρ and a corresponding sample estimator computed during the course of simulation. Experiments with simulations of the M/M/c queue with $c = 1,2,4$ and $\rho = .7,.8,.9,.95$ were conducted to evaluate the rule. The experiments used a starting rule (Rule 1) proposed in Adlakha and Fishman (1979) to reduce bias due to the initial conditions and also used the autoregressive method to obtain interval estimates of the steady-state mean. For $\rho = .7, .8, .9,$ the coverage rates are close to the specified theoretical coverage rates and are higher than those reported in the literature for other methods of interval estimation. The data reveal a degradation in the coverage rate for increasing values of activity level. For $\rho = .95$ the coverage rates are somewhat lower than those expected theoretically, indicating room for some improvement in technique. The sample sizes used to obtain the coverage rates are moderate and are insensitive to variation in the number of servers and the activity level. The rule can be easily generalized to a wider class of queueing simulations. Furthermore, experiments with a fixed truncation starting rule and a fixed sample size stopping rule clearly demonstrate the superiority of using Rule 1 and Rule 2 together. This is very encouraging, for it indicates

a procedure now exists for controlling the detrimental effects of initial conditions and skewness on interval estimation in queueing simulations.

ACCESSION for
NTIS White Section
DDC Buff Section
UNANNOUNCED
JUSTIFICATION
BY
DISTRIBUTION/AVAILABILITY CODES
Dist. AVAIL. and/or SPECIAL
A

1. Introduction

This paper proposes a stopping rule for terminating data collection in a *queueing* simulation experiment. Data collected in this way are used to compute interval estimates (for the population mean) with coverage rates that compare favorably with theoretically specified rates. This comparability is the appeal of the rule. Although most analysts of output from discrete event simulations with stochastic input would agree that interval estimates of parameters of interest are useful, prior attempts to develop procedures for computing such *useful* estimates have met with limited success. The principal impediment has been the inability of proposed techniques to produce empirical coverage rates as high as the theoretically specified coverage rates in sampling experiments designed to evaluate the proposed estimation techniques. A review of past performance appears in Adlakha (1979).

The reasons for these inadequacies are not difficult to find. Let $\hat{\theta}$ be a point estimate of θ and let $\hat{\mathrm{var}}(\hat{\theta})$ be a point estimate of $\mathrm{var}(\hat{\theta})$. Then a $1-\alpha$ interval estimate of θ usually takes the form $\hat{\theta} \pm Q\sqrt{\hat{\mathrm{var}}(\hat{\theta})}$, Q being the $1-\alpha/2$ quantile of either the normal or a t distribution. Now the writers, as well as others, have observed two potential sources of difficulty with this interval estimate. In particular, one customarily treats $\hat{\theta}$ as normally distributed when it often exhibits positive skewness in practice. Also, $\mathrm{cov}[\hat{\theta}, \hat{\mathrm{var}}(\hat{\theta})] > 0$, implying that below (above) average $\hat{\theta}$ occurs with below (above) average $\hat{\mathrm{var}}(\hat{\theta})$. These observations lead to the intuitive conclusions: Positive skewness and positive correlation imply that $\hat{\theta}$ falls below θ more than fifty percent of the time and the interval estimate width $2Q\sqrt{\hat{\mathrm{var}}(\hat{\theta})}$ is shorter

than normal theory suggests at least fifty percent of the time. This asymmetry of the center of the interval estimate and its shorter-than-expected width lead to a lower coverage rate in practice than one would expect for a normally distributed $\hat{\theta}$ and uncorrelated $\hat{\theta}$ and $\hat{\text{var}}(\hat{\theta})$.

To overcome the inadequacy of past suggestions, this paper presents an approach to interval estimation that uses ancillary information available to an analyst before, during and after execution of a simulation of a queueing system. This information consists of arrival and service rates or their corresponding reciprocals, mean interarrival and mean service times. The principal idea is to compute sample quantities such as sample mean interarrival and service times and compare them with their corresponding true values. When the deviations between sample and corresponding true quantities fall within specified tolerances, one stops or terminates the run and uses the data collected so far to compute an interval estimate. The rationale here is that making sample quantities representative of their corresponding true quantities may make $\hat{\theta}$ representative of θ . If this behavior materializes, then one intuitively expects greater success in computing adequate interval estimates for θ . One additional motivation for this use of *a priori* information is that it is available for a wide variety of queueing simulations.

The paper recommends a *stopping rule* that exploits this relationship between the theoretical and sample activity levels in a queueing simulation. When used with a *starting rule* described in Adlakha and Fishman (1979) for beginning data collection, and with the autoregressive method of interval estimation (Fishman 1971), the proposed stopping rule leads to coverage rates comparable to theoretically specified ones for a wide range of activity levels. Also, the coverage rates are higher than those reported in the

literature for other methods of interval estimation and require smaller sample sizes (for comparison see Law 1977, Fishman 1978a). We demonstrate this success by application of our proposed starting rule - stopping rule procedure to simulations of the M/M/1, M/M/2 and M/M/4 queueing simulations with activity levels ρ = .7, .8, .9 and .95 . Section 2 formulates the problem to be investigated, describes the proposed stopping rule and how it is to be used with the Starting Rule 1 in Adlakha and Fishman (1979). Section 3 presents results on coverage rates and related performance measures, showing the favorable performance of the proposed starting rule - stopping rule. It also includes an analysis of sampling variation in stopping times when using the proposed stopping rule. Section 4 compares the proposed starting rule - stopping rule procedure with several potentially useful alternatives and demonstrates the superior performance of our proposal.

2. Problem Formulation

Consider a simulation model of a queueing system with c servers in parallel, independent interarrival times with mean $1/\lambda$ and independent service times with mean $1/\omega$. Let T_i denote the elapsed time between arrivals of jobs $i-1$ and i and let S_i denote the service time of arrival i . Assume that the simulation begins with the arrival of job 1 to an empty queue and c idle servers. Let X_i denote the *system time* of completion i where system time denotes waiting time plus service time. Assume that an ultimate objective of analysis is to compute an interval estimate of the mean system time from a sample record of system times. Also assume that given the choice between starting data collection in an undercongested or a congested system, one prefers the congested one. See Adlakha and Fishman (1979) for a discussion of this point.

Starting Data Collection

After n completions occur during a simulation run, one can estimate $1/\lambda$ and $1/\omega$ by $n^{-1}\sum_{i=1}^{n} T_i$ and $n^{-1}\sum_{i=1}^{n} S_i$ respectively. These estimates are unbiased and independent of initial conditions. Now the *activity level* or traffic intensity for a queueing system usually is defined as

$$\rho = \text{arrival rate/ number of servers} \times \text{service rate}$$
$$= \lambda/(c \cdot \omega)$$

for which one estimate is

$$\tilde{\rho}_n = \sum_{i=1}^{n} S_i/(c \cdot \sum_{i=1}^{n} T_i) \ .$$

Since $\tilde{\rho}_n$ is usually a biased estimator of ρ an alternative, presumably more desirable, estimator is

$$\hat{\rho}_n = \frac{\rho\,\tilde{\rho}_n}{E(\tilde{\rho}_n)} \ , \tag{1}$$

since $E(\tilde{\rho}_n)$ is in principle derivable for most common interrarival and service time distributions. For example, in the case of exponential interarrival and service times $E(\tilde{\rho}_n) = \rho n/(n-1)$ so that

$$\hat{\rho}_n = (n-1)\tilde{\rho}_n/n \ .$$

In the present paper, data collection begins with system time $T+1$ where

Starting Rule 1: $T = \min(n: S_1^*, S_2^*$ and S_3^* hold)

and

$$S_1^* = \{|\hat{\rho}_{I,n} - \rho| \le \delta\} ,$$

$$S_2^* = \{\hat{\rho}_{I,n} > \hat{\rho}_{I,n-1}\} ,$$

$$S_3^* = \{mI \ne n-1\} ,$$

δ = specified tolerance > 0 , m = specified integer > 0 , $I = \lfloor (n-1)/m \rfloor$.

The quantity $\hat{\rho}_{I,n}$ is the *local* sample activity level,

$$\hat{\rho}_{I,n} = \frac{\sum_{i=mI+1}^{n} S_i}{c \cdot \sum_{i=mI+1}^{n} T_i} \cdot \frac{n-mI-1}{n-mI} .$$

This sample activity level applies for the M/M/c queue and requires an adjustment in its correction factor $(n-mI-1)/(n-mI)$ for distributions other than the exponential. Adlakha and Fishman (1979) describe the detailed benefits of this rule.

In words, Rule 1 requires one to use a sample activity level based on at most m past completions. The quantity I denotes the number of times one needs to reset $\hat{\rho}_{I,n}$; *i.e.*, the number of iterations minus one. A little thought shows that

$$pr(I = i) = (1 - q_m)^i \, q_m \qquad i = 0,1,...$$

where q_m is the probability of success on a given iteration. Then I has a geometric distribution with mean $(1 - q_m)/q_m$ and variance $(1 - q_m)/q_m^2$. Also, the mean number of completions $E(T)$ required to

meet Starting Rule 1 satisfies

$$m(1 - q_m)/q_m < E(T) \le m/q_m .$$

Following the Adlakha and Fishman (1979) recommendations, we choose $m = 5000$ and $\delta = .0001$.

Selecting a Stopping Rule

The stopping rule that we develop is based on a comparison of the theoretical activity level ρ and the sample activity level $\hat{\rho}_n$ in (1) computed during the execution of a simulation and where n now counts the number of completions in the collected data. As stated earlier, we anticipate that representing ρ by $\hat{\rho}_n$ will induce an underlying correlation structure in the data that leads to a representative sample mean $\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_{T+i}$, where $X_{T+1}, \ldots, X_{T+n}$ are the collected system times and will produce an interval estimate or confidence interval whose coverage rate agrees with the corresponding theoretical coverage rate.

Although the notion of a stopping rule based on a comparison of sample and corresponding theoretical activity levels is easily appreciated, several issues play an important role in determining the explicit form a stopping rule should assume. These issues are:

i. What should the tolerance be for the comparison of $\hat{\rho}_n$ and ρ ?

ii. Should there be a *directional* relationship of $\hat{\rho}_n$ to $\hat{\rho}_{n-1}$?

iii. Should there be a lower bound on the size of the sample to be obtained for estimation?

iv. Should there be an upper bound on the sample size to avoid excessive data collection?

We address each issue in the context of the formulated stopping procedure:

1. Begin the simulation in the empty and idle state.
2. Apply Starting Rule 1 to determine the first system time X_{T+1} to collect.
3. Collect system times $X_{T+1}, \ldots, X_{T+N}$ where

$$N = \min(n: |\hat{\rho}_n - \rho| \le \delta^*)$$

and $\hat{\rho}_n$ is based on interarrival and service times for completions $T+1, \ldots, T+n$ and δ^* is a specified tolerance.

Stopping Rule 2 (steps 3 above)

4. Compute a sample mean and interval estimate from $X_{T+1}, \ldots, X_{T+N}$.

Tolerance Criterion

Step 3 implies that at the termination of data collection $\hat{\rho}_N \in [\rho-\delta^*, \rho+\delta^*]$. For values of $\hat{\rho}_N \in [\rho-\delta^*, \rho)$ it is conceivable that $\hat{\rho}_n < \rho-\delta^*$ for all $n < N$ and that a congested state is never experienced. Alternatively, when

$$\rho \le \hat{\rho}_N \le \rho + \delta^*$$

it is possible that $\hat{\rho}_n > \rho + \delta^*$ for all $n < N$ and that the system fails to experience less than average congestion. Moreover, the ratio $\hat{\rho}_n/\rho$ has a positively skewed distribution with mean 1 and is independent of ρ for a fixed δ^* (see Adlakha 1979). This indicates that as

ρ increases $pr(\hat{\rho}_n < \rho - \delta^*)$ increases faster than $pr(\hat{\rho}_n > \rho + \delta^*)$, implying that it is more likely that $\hat{\rho}_N \in [\rho-\delta^*,\rho)$ than $\hat{\rho}_N \in [\rho, \rho+\delta^*]$.

It is reasonable to believe that if the choice lies between over-estimating or underestimating the congestion, most analysts would prefer to overestimate in queueing models. This is especially important in the case of systems with high activity level ρ . To favor this behavior in the system, we modify step 3 by the substitution:

$$N = \min\{n: \rho \le \hat{\rho}_n \le \rho + \delta^*\} .$$

Directional Criterion

Once $\hat{\rho}_n$ is within the specified interval $[\rho, \rho+\delta^*]$, it indicates that the desired congestion level has been reached. At this stage it would be preferable to continue collecting data and stop only when there is an indication that the system is leaving the congested state, *i.e.*, $\hat{\rho}_n < \hat{\rho}_{n-1}$. We expect this to happen when completion n has a shorter service time or a longer interarrival time than the average, thereby contributing to a below average system time for the nth customer. This will also produce a decrease in $\hat{\rho}_n$. Therefore, we include an additional condition "$\hat{\rho}_n < \hat{\rho}_{n-1}$" in step 3 . We anticipate that this condition along with the condition "$\rho \le \hat{\rho}_n \le \rho + \delta^*$" will provide a sample record with enough observations from the congested system.

Lower Bound on the Sample Size

A precise estimate of $\mathrm{var}(\bar{X}_n)$ is important for interval estimation. As mentioned earlier, we use the autoregressive analysis approach to estimate $\mathrm{var}(\bar{X}_n)$. Based on experiments, Fishman (1971) showed that this sample variance approximation is not reasonable when the sample size $n < 500$. To improve on this approximation, we restrict stopping time to $n \geq 1000$ in step 3.

Upper Bound on Sample Size

Experience in testing starting rules in earlier work (Fishman and Moore 1978) revealed that in the absence of an iterative procedure as in Starting Rule 1 excessively large starting times can occur. Step 3 admits a similar possibility for stopping times, as preliminary experimentation confirmed. To favor congested systems and to avoid excessively large stopping times, we propose an iterative stopping rule with upper bound of $m^* = 5000$, to be consistent with Starting Rule 1. The upper bound is applied as follows: If the stopping rule is not satisfied until $n = 5000$ during a simulation run, these observations are discarded. Another independent replication is generated from the empty and idle state and Starting Rule 1 is again satisfied. Then the complete stopping rule is:

$$N = \min(n\colon\ \rho \leq \hat{\rho}_n \leq \rho+\delta^*,\ \hat{\rho}_n < \hat{\rho}_{n-1},\ n \geq 1000,\ m^* = 5000)\ .$$

A perusal of the data obtained from the initial experiments indicates that for replications where $\hat{\rho}_{N-1} > \rho+\delta^*$, the empirical coverage rates are comparable to the theoretically specified coverage rates. A little thought shows that this is not surprising. It is reasonable to believe that the value of $\hat{\rho}_N$ might play a role in determining the

coverage rate. The directional criterion $\hat{\rho}_N > \hat{\rho}_{N-1}$ does not appear to induce sufficient congestion in the system although $\rho \le \hat{\rho}_N \le \rho+\delta^*$, because it is possible that $\hat{\rho}_n < \rho$ for $n \le N-2$. In contrast, the condition $\hat{\rho}_{N-1} > \rho+\delta^*$ favors more congestion in the system.

The proposed procedure with lower and upper bounds is now:

1. Begin the simulation in the empty and idle state.
2. Apply Starting Rule 1 to determine the first system time X_{T+1} to collect.
3. $n \leftarrow 0$.
4. $n \leftarrow n + 1$.
5. Collect system time X_{T+n}.

Stopping Rule 2

6. If $n \le 1000$ go to 4.
7. Compute $\hat{\rho}_n$ based on interarrival and system times for completions $T+1, \ldots, T+n$.
8. If $\rho \le \hat{\rho}_n \le \rho + \delta^*$ and $\hat{\rho}_{n-1} > \rho + \delta^*$, $N \leftarrow n$ and compute a sample mean and interval estimate from $X_{T+1}, \ldots, X_{T+N}$ and stop.
9. If $n = m^* = 5000$, discard $X_{T+1}, \ldots, X_{T+n}$ and go to step 1.

10. Go to step 4.

Although we anticipate that the upper bound m^* would keep the sample size from becoming excessive, discarding the data on a replication when the stopping rule is not satisfied may affect the statistical reliability of the estimate. We study the effect of this upper bound in Section 3. Preliminary testing indicated that $\delta^* = .01$ would be a worthwhile choice.

We study Stopping Rule 2 using the experimental design:

$$\rho = .7,\ .8,\ .9,\ .95\ ,$$
$$c = 1,\ 2,\ 4\ .$$

The objective here is to investigate the performance of Stopping Rule 2 for varying number of servers in the M/M/c queue over a range of values of ρ that represent moderate to high levels of congestion.

3. Results of Experimentation

To study Stopping Rule 2, we performed 100 independent replications on each queueing model $c = 1, 2, 4$ for each activity level ($\rho = .7, .8, .9, .95$). On each replication, the collected data were:

N^* = stopping time, including the number of observations in discarded replications,

$\bar{X}_N$ = sample mean system time based on $N = N^*$ (mod 5000) observations,

$\hat{\rho}_N$ = sample activity level at termination

R_j = sample autocovariance at lag j, $j = 0,1,\ldots,50$.

The sample mean and autocovariances were used with the autoregressive

method to produce an estimate of $\mathrm{var}(\bar{X}_N)$ and 90 and 95 percent interval estimates for mean system time.

Coverage Analysis

Table 1 presents the empirical coverage rates for mean system time. The data show that the coverage rates obtained for $\rho \le .9$ are comparable to those suggested by the theory. Notice that for a given value of c, the coverage rate tends to decrease as ρ increases. This implies that higher levels of utilization in the system degrade the performance of Stopping Rule 2. For a given value of ρ, the performance appears to be unaffected by different values of c.

Since, to our knowledge, no results on the estimation of the mean system time μ for the M/M/c queue with $c > 1$ and $\rho = .95$ have been reported in the literature, we use the coverage rate for the M/M/1 queue with $\rho = .9$ obtained from earlier proposed methods as our basis of comparison. For a sequential method developed in Fishman (1971) the 90 percent empirical coverage rates obtained were between 66 percent and 79 percent. In an empirical testing of the batch means method, Law (1977) obtained a 90 percent empirical coverage rate of 86 for a fixed sample size of 12800. With the method of batch means, Fishman (1978a) obtained an empirical coverage rate of 90 for the 95 percent confidence intervals with a fixed sample size of 16384. Therefore, we can conclude that Stopping Rule 2 provides a better coverage rate than the methods cited in the aforementioned references. The significance of this achievement is magnified by the fact that the sample sizes used to obtain these coverage rates are much smaller than those used by other methods. We discuss the distribution of the sample size shortly.

Table 1
Empirical Coverage Rates Using
Stopping Rule 2 for Mean System Time

c \ ρ	90% Coverage .7	.8	.9	.95	95% Coverage .7	.8	.9	.95
1	97	91	88	77	99	94	93	82
2	97	96	90	73	97	97	92	79
4	95	95	86	73	98	97	90	83

Bias

In general one would like sample means that are unbiased estimators of mean system time μ. However, because of the procedures for starting and stopping data collection, one has reason to believe *a priori* that $\bar{X}_N$ for random T and N is a biased estimator of μ.

Let Y_i denote the sample mean of replication i for $i = 1, \ldots, 100$. Then

$$\bar{Y} = \frac{1}{100} \sum_{i=1}^{100} Y_i$$

gives the grand sample mean over all replications for a given (c, ρ) 2-tuple and $\bar{Y} - \mu$ gives an estimate of the bias in $Y_1, \ldots, Y_{100}$.

Then

$$s^2_{\overline{Y}} = \frac{1}{100\times 99} \sum_{i=1}^{100} (Y_i - \overline{Y})^2$$

is an estimate of $\mathrm{var}(\overline{Y})$.

Table 2 shows the relative bias $(\overline{Y} - \mu)/\mu$ and Table 3 shows the standardized deviate $(\overline{Y} - \mu)/\sqrt{s^2_{\overline{Y}}}$. Although the significance of the bias is well established in Table 3, its small relative impact in Table 2 removes any substantive concern about an overriding bias component.

Table 2

Relative Bias of the Sample Mean System Time[a]

$(\overline{Y} - \mu)/\mu$

c \ ρ	.7	.8	.9	.95
1	.046	.044	.057	.077
2	.030	.048	.076	-.051
4	.036	.034	.036	.056

[a] The mean system time μ can be computed using formulae in Gross and Harris (1974).

Table 3

Standardized Bias of the Sample Mean System Time

$(\overline{Y} - \mu)/\sqrt{s^2_{\overline{Y}}}$

c \ ρ	.7	.8	.9	.95
1	4.409*	2.792*	1.874	1.314
2	3.605*	3.609*	2.721*	-1.287
4	4.879*	3.243*	1.411	1.092

*Significant at the 5% level.

Variance Analysis

When computing confidence intervals for $\overline{X}_N$, one relies upon the accuracy of the estimate of its variance. Here we examine the issue of bias in the estimation of the variance of the sample mean as a possible source of error.† As previously discussed, the autoregressive procedure was employed to estimate the variance of the sample mean in a replication. Let Y_i and $\hat{V}_i$ denote the sample mean and the sample variance of the sample mean, respectively, on the ith replication. Furthermore, let

$$\hat{V} = \frac{1}{100} \sum_{i=1}^{100} \hat{V}_i ,$$

†Additional comparisons appear in Adlakha (1979).

and

$$\tilde{V} = \frac{1}{99} \sum_{i=1}^{100} (Y_i - \bar{Y})^2 ,$$

where

$$\bar{Y} = \frac{1}{100} \sum_{i=1}^{100} Y_i .$$

For $\hat{V}_i$ to be an adequate estimator of $\mathrm{var}(Y_i)$, $\hat{V}$ should be close to $\tilde{V}$ since $\tilde{V}$ is an unbiased estimator of $\mathrm{var}(Y_i)$.

Table 4 presents the variance estimates $\hat{V}$ and $\tilde{V}$. The results show an upward bias in $\hat{V}$ for $\rho \le .9$. This pattern is somewhat different for $\rho = .95$, where in two out of three cases a downward bias in $\hat{V}$ is suggested. No doubt, this underestimate for $\rho = .95$ is responsible for the degradation in coverage rate in Table 1. Clearly, work remains to be done in improving $\hat{V}$ as an estimator of $\mathrm{var}(\bar{X}_N)$.

Correlation

Fishman (1978a) observed that a principal source of error in the interval estimation of correlated data is the high correlation between $\bar{X}_N$ and $\hat{\mathrm{var}}(\bar{X}_N)$. Table 5 provides these sample correlations for the experiments in this study. The data confirm that $\bar{X}_N$ and $\hat{\mathrm{var}}(\bar{X}_N)$ are positively correlated. At first thought, this observation is perplexing, since we have pointed to correlation's being the culprit that degrades the coverage rate. One plausible explanation lies in the

Table 4

Comparison of Estimates of $var(\overline{X}_N)$

c \ ρ		.7	.8	.9	.95
	$\hat{V}$	.107	.580	13.747	64.700
1	$\tilde{V}$	.060	.393	7.380	124.557
	Ratio	1.78	1.48	1.86	0.51
	$\hat{V}$	.099	.699	11.292	61.629
2	$\tilde{V}$	.053	.345	7.060	59.568
	Ratio	1.87	2.03	1.60	1.03
	$\hat{V}$	.121	.679	13.383	69.486
4	$\tilde{V}$	.080	.336	7.386	111.770
	Ratio	1.51	2.02	1.81	0.62

Table 5

Sample Correlation[a] between $\overline{X}_N$ and $\widehat{var}(\overline{X}_N)$

c \ ρ	.7	.8	.9	.95
1	.773	.803	.802	.700
2	.781	.767	.671	.541
4	.794	.742	.702	.570

[a]The critical value with 100 observations at the 5% significance level is .195 .

significant bias in $\overline{X}_N$ (see Table 3). This small but perceptible upward tendency may be sufficient to make the correlation in Table 5 work in our favor. That is a tendency to overestimate μ with a positive correlation between $\overline{X}_N$ and $\hat{var}(\overline{X}_N)$ leads to wider interval estimates than strict normal theory would suggest. In turn, this may contribute to the improved coverage rates. In fact, the absence of significant bias in $\overline{X}_N$ for $\rho = .95$ may be the missing ingredient needed to make $\hat{V}$ more representative of $var(\overline{X}_N)$.

Recall that Stopping Rule 2 imposes an upper bound m^* on the sample size in a replication. To study the affect of this bound on $\overline{X}_N$, we examine the sample correlations between the stopping time N^* and the corresponding sample mean $\overline{X}_N$, $N = N^*(\text{mod } 5000)$ in Table 6 . The correlations show no significance at the 5 percent level.

Table 6

Sample Correlation[a] Between $\overline{X}_N$ and N^*

c \ ρ	.7	.8	.9	.95
1	-.049	.043	.047	-.100
2	.087	.106	.192	.068
4	-.049	-.007	.015	.093

[a] The critical value with 100 observations at the 5% significance level is .195 .

Distribution of Stopping Time

Table 7 gives the sample quantiles of the stopping time N^* . The lower quantile values reflect the lower bound of 1000 in Stopping Rule 2 . Although the quantiles show the distribution of the stopping time to be positively skewed, the 95 percent quantiles range from 8722 to 17156 , indicating small probability for an excessive sample size.

Notice that the quantiles appear to be insensitive to ρ and c . Since the distribution of $\hat{\rho}_n/\rho$ is independent of c , the insensitivity of N^* to c does not come as a surprise. The insensitivity to ρ is no doubt partially due to the facts that for a fixed δ^* and n , as ρ increases the $\mathrm{pr}(\rho \le \rho_n \le \rho + \delta^*)$ decreases, whereas $\mathrm{pr}(\hat{\rho}_{n-1} > \rho + \delta^*)$ increases. The effects of these conditions on the stopping time seem to balance each other for an increase in ρ .

The sample mean, standard deviation and coefficient of variation of the stopping time N^* appear in Table 7 . The mean stopping times vary between 3586 and 5027 and occur between the 60 and 70 percentiles. The sample coefficients of variation $\hat{\nu}_{N^*}$ are generally close to one. These observations suggest that although the distribution of N^* is positively skewed, it has a short tail to the right. Thus, encountering an excessive stopping time is a remote possibility.

4. Comparison with Other Rules

This section compares the performance of the proposed starting-stopping rule procedure with some alternative starting and stopping rules using the experimental design which we describe shortly. We use the coverage rate of

Table 7

Sample Quantiles q of Stopping Time N* for Stopping Rule[a] 2

$Pr(N^* \le q) = p$

c	1				2				4			
100×p \ p	.7	.8	.9	.95	.7	.8	.9	.95	.7	.8	.9	.95
1	1001	1000	1000	1000	1000	1000	1000	1007	1002	1002	1000	1000
2	1003	1001	1000	1001	1001	1009	1002	1007	1002	1003	1001	1002
5	1036	1007	1001	1009	1011	1011	1010	1017	1018	1007	1007	1009
10	1085	1043	1004	1036	1070	1041	1036	1035	1052	1027	1042	1031
15	1095	1077	1009	1048	1122	1075	1086	1054	1103	1066	1062	1102
20	1170	1115	1036	1103	1129	1118	1163	1067	1179	1151	1089	1144
25	1268	1186	1086	1131	1350	1174	1227	1091	1235	1317	1144	1195
30	1351	1259	1255	1199	1658	1274	1332	1184	1313	1506	1179	1335
35	1428	1398	1322	1383	1760	1399	1476	1332	1567	1831	1352	1368
40	1639	1589	1459	1469	2066	1666	1550	1552	1720	2141	1496	1481
45	1934	1797	1530	1596	2361	1797	1752	1872	1916	2533	1657	1777
50	2048	2048	1724	1755	3053	2038	1930	2199	2095	2910	1858	2115
55	2403	2384	1839	2134	3732	2197	2152	2727	2551	3480	2074	2604
60	3167	2921	2107	2495	3987	2603	3006	4130	3345	4119	2705	3208
65	4382	4035	2677	3385	6005	3297	4789	6077	4348	6012	2868	4108
70	6038	4979	3180	6017	6279	4510	6134	6316	6021	6065	6003	6003
75	6215	6328	3972	6147	6789	6022	6573	6646	6116	6465	6060	6077
80	6789	6787	6101	6255	8279	6385	7219	7423	6231	6896	6265	6233
85	8105	7226	6387	6705	11078	7340	11014	9088	6343	8487	6553	6678
90	11513	11209	7678	7126	11281	11033	11727	11089	7712	11274	7655	7638
95	17156	13358	11003	8722	16056	11411	14995	11848	12829	16073	12151	11528
98	26071	16443	11481	11202	16654	16773	16253	13169	21204	17096	21491	16082
99	31011	17839	11902	11421	17198	17625	16393	14027	22494	17344	22183	17114
min	1001	1000	1000	1000	1000	1000	1000	1007	1002	1002	1000	1000
max	37107	21470	16716	23160	21424	46851	26046	37045	23025	17575	36016	21437
$\bar{N}^*$	5018	4188	4453	3586	[illegible]	4228	4652	4627	4295	4757	4128	4026
$\hat{\sigma}_{N^*}$	6378	4330	4759	3446	4647	5713	4913	4989	4627	4376	5337	3996
$\hat{\nu}_{N^*}$	1.27	1.03	1.07	.96	.92	1.35	1.06	1.08	1.08	.92	1.29	.99

[a]The quantities $\hat{\sigma}_{N^*}$ and $\hat{\nu}_{N^*}$ denote the standard deviation and the coefficient of variation respectively.

the mean system time as the criterion for comparisons.

Let R_1 and R_2 denote Starting Rule 1 and Stopping Rule 2 respectively. As a variation of rule R_1, consider a fixed truncation rule that puts the M/M/c queue into the steady state and denote this rule as F_T. Similarly, as an alternative for rule R_2, consider a fixed sample size rule and denote it as F_S. Rule F_S uses a sample size equivalent to the mean stopping time of rule R_2 for given c and ρ.

Four combinations of starting and stopping rules are considered:

i. $(R_1, R_2) \equiv$ Starting Rule 1 with Stopping Rule 2 ,

ii. $(F_T, R_2) \equiv$ fixed truncation with Stopping Rule 2 ,

iii. $(R_1, F_S) \equiv$ Starting Rule 1 with fixed sample size,

iv. $(F_T, F_S) \equiv$ fixed truncation with fixed sample size.

Hereafter, allusion to *system* (A,B) means the experiment with starting rule A and stopping rule B. System $(\cdot, B)$ refers to stopping rule B where starting rule can be either R_1 or F_T and system $(A, \cdot)$ refers to starting rule A with stopping rule R_2 or F_S.

Let (F_0, F_S) denote the system where simulation begins in the empty and idle state and F_S is the stopping rule as defined earlier. Blomqvist (1970) showed that, provided certain conditions are satisfied, for a large sample size the mean-square error of sample mean as an estimator of the population mean is minimized if one starts the simulation from the empty and idle state and no observations are discarded. However, for a correlated sample record, as is the case in most queueing simulations, it is not clear that minimizing mean-square error would necessarily provide the theoretically specified coverage rate. Therefore, we also

compare the mean-square errors and coverage rates obtained by using system (R_1, R_2) with those obtained from system (F_0, F_S) .

To study the performance of the aforementioned systems, we consider the experimental design:

$$\rho = .7, .8, .9, .95 ,$$

$$c = 1, 2, 4,$$

$$\text{system} = (R_1, R_2), (F_T, R_2), (R_1, F_S), (F_T, F_S) .$$

With regard to truncation rule F_T , an exploratory research showed that a truncation of 1000 observations, starting from the empty and idle initial state, suffices to put a system in the steady state for the M/M/c queue with $c = 1, 2, 4$ and $\rho = .7, .8, .9$. This implies that the initial conditions hardly affect completion 1001 for these queues. However, for $\rho = .95$, a truncation of 1000 observations was not sufficient, but a truncation of 2000 observations appeared adequate.

With regard to stopping rule F_S of fixed sample size, for each given value of ρ and c we take a sample of size $\overline{N}^*$, the corresponding mean stopping time obtained from rule R_2 . This makes rules R_2 and F_S comparable in the sense that the mean number of observations are identical.

Table 8 presents the results for 100 replications for each value of ρ and c with systems (F_T, R_2), (R_1, F_S), (F_T, F_S) and the original results for (R_1, R_2) . All results are based on the autoregressive analysis in Fishman (1978b). Three tendencies are apparent. Firstly, the degradation in the coverage rate for all the systems as ρ increases.

Table 8

Coverage Rates for Alternative Systems

		90% Coverage				95% Coverage			
		ρ				ρ			
c	system	.7	.8	.9	.95	.7	.8	.9	.95
1	(R_1, R_2)	97	91	88	77	99	94	93	82
	(F_T, R_2)	97	94	87	80	97	95	88	84
	(R_1, F_S)	88	72	75	63	94	82	81	69
	(F_T, F_S)	86	75	83	73	90	81	87	78
2	(R_1, R_2)	97	96	90	73	97	97	92	79
	(F_T, R_2)	92	96	87	68	95	97	90	80
	(R_1, F_S)	84	73	69	63	89	81	77	67
	(F_T, F_S)	89	78	82	66	90	83	84	76
4	(R_1, R_2)	95	95	86	73	98	97	90	83
	(F_T, R_2)	98	98	86	75	99	98	88	80
	(R_1, F_S)	82	83	78	78	91	86	82	82
	(F_T, F_S)	84	87	77	70	86	88	82	76

This shows that high utilization in the M/M/c queue affects the performance uniformly for each system. Secondly, the superiority in the performance of system (R_1, R_2) over (R_1, F_S) and of system (F_T, R_2) over (F_T, F_S). This indicates that, regardless of the starting criterion used, stopping rule R_2 yields better coverage rates than the fixed sample size F_S with equivalent expected cost, in terms of number of observations. Thirdly, there is no apparent distinction in the performance of systems (R_1, R_2) and (F_T, R_2), and of systems (R_1, F_S) and (F_T, F_S).

This implies that the performance of Starting Rule 1 is as good as the performance of fixed truncation which puts the system in the steady state.

Two subsequent statistical analyses also revealed useful insights.[†] An analysis of variance together with multiple comparison procedures based on transformed coverage rates showed that for $\rho \le .9$, system $(\cdot, R_2)$ was significantly better than system $(\cdot, F_S)$. Moreover, no significant difference existed between the performance of systems $(R_1, \cdot)$ and $(F_T, \cdot)$. To select the best among the four systems, the procedure in Bechhofer (1954) was used. It indicated that system (R_1, R_2) was best for $\rho = .7$ and $.9$, with at least probabilities of .562 and .956 respectively. For $\rho = .8$, (F_T, R_2) was selected as best with a correct selection probability of .458. For $\rho = .95$, the procedure was indifferent between systems (R_1, R_2) and (F_T, R_2). The prevalence with which the several statistical analyses focus on R_1 and R_2 enables us to conclude that (R_1, R_2) is at least as good in general as any of the other systems. However, the greatest attraction of (R_1, R_2) comes from the fact that F_T is impractical in application unless one knows how many observations to truncate and F_S is impractical unless one knows how many observations to collect to make the autoregressive analysis a useful tool.

As mentioned earlier, the Blomqvist (1970) result encourages one to truncate no observations to minimize mean-square error. To test the performance of the zero truncation rule F_0, we ran the design $\rho = .7, .8, .9$ and $.95$ and $c = 1, 2$ and 4 for system (F_0, F_S) and computed coverage rates. Table 10 compares these rates with those for

---

[†] See Adlakha (1979) for details.

(R_1, R_2). The superiority of (R_1, R_2) is clearly established, in spite of the fact that computed mean-square errors for this system were larger than corresponding mean-square errors for (F_0, F_S). These observations lead to the conclusion that (F_0, F_S) generates a more accurate point estimate in terms of mean-square error but (R_1, R_2) gives a superior interval estimate as measured by coverage rate. In our opinion the greater precision of (F_0, F_S) for the point estimate is of little value if we have no adequate way to estimate the precision. We prefer to pay the penalty of a less accurate point estimate whose precision we can successfully evaluate.

Table 9

Coverage Rates for Systems (F_0, F_S) and (R_1, R_2)

		90% Coverage				95% Coverage			
			ρ				ρ		
c	system	.7	.8	.9	.95	.7	.8	.9	.95
1	(F_0, F_S)	79	82	76	57	83	89	83	60
	(R_1, R_2)	97	91	88	77	99	94	93	82
2	(F_0, F_S)	89	78	71	68	93	85	78	74
	(R_1, R_2)	97	96	90	73	97	97	92	79
4	(F_0, F_S)	90	82	74	62	94	85	81	68
	(R_1, R_2)	95	95	86	73	98	97	90	83

5. Conclusions

The objective in this study has been to devise generally applicable methods to improve interval estimation in a queueing simulation. On the basis of accumulated empirical evidence presented in this paper, we continue to recommend

Rule 1:

$$T = \min\{n: |\hat{\rho}_{I,n} - \rho| \le \delta,\ \hat{\rho}_{I,n} > \hat{\rho}_{I,n-1},\ mI \ne n-1,\ \delta = .0001,\ m = 5000\}$$

as a starting rule (see Adlakha and Fishman 1979). Also, we recommend

Rule 2:

$$N = \min\{n: \rho \le \hat{\rho}_n \le \rho+\delta^*,\ \hat{\rho}_{n-1} > \rho+\delta^*,\ n \ge 1000,\ m^* = 5000,\ \delta^* = .01\}$$

as a stopping rule.

The rules give satisfactory performance for multiserver queues with moderate-to-high levels of congestion. Both rules use a moderate number of observations, are easy to understand and to implement. Furthermore, experiments with a fixed truncation starting rule and a fixed sample size stopping rule clearly demonstrate the superiority of using Rule 1 and Rule 2 together. This is very encouraging, for it indicates a procedure now exists for controlling the detrimental effects of initial conditions and skewness on interval estimation in queueing simulations.

6. References

Adlakha, Veena G. (1979). "Starting and Stopping Rules for Data Collection in Discrete-Event Simulation," Unpublished Ph.D. Dissertation, Curriculum in Operations Research and Systems Analysis, The University of North Carolina at Chapel Hill.

Adlakha, Veena G. and George S. Fishman (1979). "A Starting Rule for Data Collection in Queueing Simulations," Technical Report 79-3, Curriculum in Operations Research and Systems Analysis, The University of North Carolina at Chapel Hill.

Bechhofer, R. E. (1954). "A Single-sample Multiple Decision Procedure for Ranking Means of Normal Populations with Known Variances." The Annals of Mathematical Statistics, 25, 16-39.

Blomqvist, N. (1967). "The Covariance Function of the M/G/1 Queueing System." Skandinivisk Aktuarietidskrift, 50, 157-74.

Fishman, George S. (1971). "Estimating Sample Size in Computer Simulation Experiments." Management Science, 18, 21-38.

________ (1978a). "Grouping Observations in Digital Simulation." Management Science, 24, 510-21.

________ (1978b). Principles of Discrete Event Simulation. New York: John Wiley and Sons.

Fishman, George S. and Louis R. Moore, III (1978). "Starting Times for Data Collection in a Queueing Simulation I: Experiments with a Single Server Model." Technical Report TR78-1, Curriculum in Operations Research and Systems Analysis, The University of North Carolina at Chapel Hill.

Gross, D. and C. Harris (1974). Fundamentals of Queueing Theory. New York: John Wiley and Sons.

Law, A. M. (1977). "Confidence Intervals in Discrete Event Simulation: A Comparison of Replication and Batch Means." Naval Research Logistics Quarterly, 24, 667-77.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 79-4	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A Stopping Rule for Data Collection in Queueing Simulations		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Veena G. Adlakha and George S. Fishman		8. CONTRACT OR GRANT NUMBER(s) N00014-76-C-0302
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of North Carolina Chapel Hill, North Carolina 27514		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Operations Research Program Office of Naval Research Arlington, Virginia 22217		12. REPORT DATE Ausust, 1979
		13. NUMBER OF PAGES 27
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this document is unlimited.

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Autoregressive method — Simulation

Interval Estimate — Stopping Rule

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper proposes a stopping rule (Rule 2) for terminating data collection in a queueing simulation experiment. The appeal of the rule lies in the fact that data collected in this way can be used to compute interval estimates with coverage rates that compare favorably with theoretically specified rates. The rule relies on a comparison between *a priori* information on the activity level (traffic intensity) ρ and a corresponding sample estimator computed during the course of simulation. Experiments with simula-

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

tions of the M/M/c queue with $c = 1,2,4$ and $\rho = .7,.8,.9,.95$ were conducted to evaluate the rule. The experiments used a starting rule (Rule 1) to reduce bias due to the initial conditions and also used the autoregressive method to obtain interval estimates of the steady-state mean. For $\rho = .7,.8,.9$, the coverage rates are close to the specified theoretical coverage rates and are higher than those reported in the literature for other methods of interval estimation. The data reveal a degradation in the coverage rate for increasing values of activity level. For $\rho = .95$ the coverage rates are somewhat lower than those expected theoretically, indicating room for some improvement in technique. The sample sizes used to obtain the coverage rates are moderate and are insensitive to variation in the number of servers and the activity level. The rule can easily be generalized to a wider class of queueing simulations. Furthermore, experiments with a fixed truncation starting rule and a fixed sample size stopping rule clearly demonstrate the superiority of using Rule 1 and Rule 2 together. This is very encouraging, for it indicates a procedure now exists for controlling the detrimental effects of initial conditions and skewness on interval estimation in queueing simulations.